Directed-Hypergraph-Based Channel Allocation for Ultradense Cloud D2D Communications With Asymmetric Interference

Youming Sun , Zhiyong Du , *Member, IEEE,* Yuhua Xu , *Member, IEEE,* Yuli Zhang, *Student Member, IEEE,* Luliang Jia , *Student Member, IEEE,* and Alagan Anpalagan , *Senior Member, IEEE*

Abstract—In this paper, we investigate channel allocation for ultradense device-to-device (D2D) communications. Different from both binary graph and undirected hypergraph interference model, an improved directed hypergraph is applied to simultaneously represent cumulative and asymmetric interference aspects in the context of ultradense communications. We formulate the channel access problem in cloud D2D communication networks as a directed-hypergraph-based local altruistic game, which is proved to be an exact potential game. Then, a multiagent concurrent learning scheme in centralized-distributed fashion is proposed to search the optimal pure Nash equilibrium, which can also maximize the normalized network capacity. Finally, simulation results are presented to validate the proposed learning scheme.

Index Terms—Directed hypergraph, Channel allocation, Device-to-device communications, Exact potential game, Learning algorithm.

I. INTRODUCTION

DEVICE-TO-DEVICE (D2D) communications has emerged as a promising technology in the fifth generation (5G) communications, because it can enable the direct transmission among communication devices without the assistance of base station (BS) [1], [2]. Accordingly, D2D communications can significantly improve the spectrum utilization efficiency from the perspective of spatial domain and alleviate wireless operator's traffic burden. In the near future, there will be numerous D2D pairs under the coverage of BS. The appropriate

channel allocation is a key issue in the resource management for wireless networks. Many graph-based allocation schemes have been proposed for D2D communications. However, most of them applied the binary interference model which only focuses on the pair-wise strong interference links. For the dense wireless networks, it cannot well model the complex interference relationship. To respond to the limitation of the binary interference graph, the hypergraph is introduced to represent the complicated interference relationship for ultra-dense networks, such as D2D networks [1] and small cell networks [4]. Hypergraph interference model can capture the cumulative interference effects caused by multiple co-channel weak interfering nodes.

There exist some studies on the hypergraph-based resource allocation for wireless networks [1]–[6]. Specifically, in [5], Li *et al.* investigated the scheduling problem in wireless ad hoc networks based on hypergraph interference model. Zhang *et al.* in [1] investigated the channel allocation for D2D underlay communications, and proposed an efficient hypergraph-coloring-based channel allocation scheme. Sun *et al.* in [2] designed a distributed learning algorithm for dense D2D communications. However, the above mentioned studies are based on undirected hypergraph and do not fully consider the asymmetric interference effect. In practical communication scenario, owing to devices' heterogenous communication capability in terms of both transmission power and receiver sensitivity, it is common that the interference relationship among communication devices tends to be asymmetric especially in hyper-densely deployment networks.

In this paper, we investigate the channel allocation for ultradense cloud D2D networks under the management of wireless operator. To be specific, we firstly present an improved directed hypergraph to represent the cumulative and asymmetric interference effects simultaneously. Then, we formulate the channel access problem in the cloud D2D networks as a local cooperation game to maximize the normalized network capacity, which is proved to be an exact potential game (EPG) admitting at least one pure Nash equilibrium (PNE). To fully utilize the cloud platform, we design a centralized-distributed learning algorithm to converge to the optimal PNE, which maximizes the defined normalized network capacity. To the best of our knowledge, this is the first work to incorporate the directed hypergraph into the resource allocation of wireless networks.

Y. Sun, Y. Xu, Y. Zhang, and L. Jia are with the Army Engineering University, Nanjing 210000, China (e-mail: sunyouming10@163.com; yuhuaenator@gmail.com; yulipkueecs08@sina.com; jiallts@163.com).

Z. Du is with the National University of Defense Technology, Changsha 410000, China (e-mail: duzhiyong2006@163.com).

A. Anpalagan is with the Department of Electrical and Computer Engineering, Ryerson University, Toronto, ON M5B 2K3 Canada (e-mail: alagan@ee.ryerson.ca).
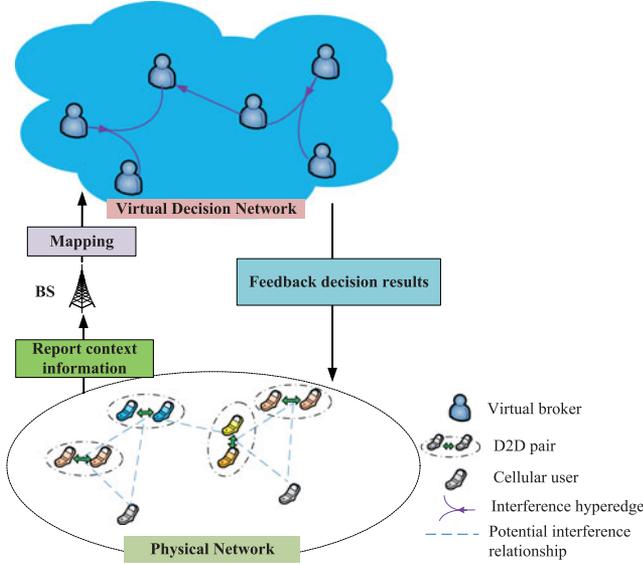
Fig. 1. Illustration of the cloud D2D network framework.

Fig. 2. An example of a undirected hyperedge with $Q = 3$, where Q is the maximum cardinality of hyperedge, and its possible directed hyperedge patterns.

II. SYSTEM MODEL AND PROBLEM FORMULATION

As shown in Fig. 1, we consider a cloud-based D2D network framework as similar as in [7], [8], where the end users, containing D2D pair and cellular user, are under the management of wireless network service provider (WSP). Initially, each user reports its individual context information including geographical location, transmission power and receiver sensitivity, etc to the nearby WSP's BS. WSP constructs the hypergraph interference relationship among the active communication nodes in current networks.[1] Then, WSP maps the physical D2D communication networks to a virtual decision network (VDN). In the VDN, WSP generates a one-to-one corresponding virtual broker (VB) for each user to make channel access decision automatically. Finally, WSP broadcasts VBs' decisions to their communication nodes in physical networks. Denote the user set, containing cellular users and D2D pairs, as $\mathbf{N} = (1, 2, \ldots, N)$, where N is total number of users. The number of licensed channels is M and assume that these channels are available for both cellular users and D2D pairs.

A. Directed Hypergraph Preliminaries

Definition 1 ([11]): **Hypergraph** $\boldsymbol{\Gamma}$ denoted by $\boldsymbol{\Gamma} = (\mathbf{V}, \mathbf{E} = (e_i)_{i \in \Lambda})$, where $\mathbf{V} = \{v_1, v_2, \ldots v_n\}$ is a finite set of vertices, is a family $(e_i)_{i \in \Lambda}$ (Λ is a finite set of indexes) of subsets of $\mathbf{V}$ and e_i is termed as hyperedge.

Hypergraph is the generalization of graph and, hyperedge can contain any subset of vertex set $\mathbf{V}$. Let n and m be the cardinality of its vertex set and hyperedge set, respectively. Generally, the hypergraph can be specified from its one-to-one corresponding $n \times m$ incidence matrix.

Definition 2: **Incidence matrix** $A = [a_{ij}]_{n \times m}$ is a Boolean matrix for a undirected hypergraph $\boldsymbol{\Gamma} = (\mathbf{V}, \mathbf{E} = (e_i)_{i \in \Lambda})$,

where $a_{ij} = 0$ or 1. If the vertex i is associated with the hyperedge j, then $a_{ij} = 1$; otherwise, $a_{ij} = 0$.

Recently, hypergraph has been introduced to represent the complex interference relationship in hyperdense communication networks [13]. Hypergraph interference model can capture cumulative interference effect resulting from multiple weak interferers, while the traditional binary graph only focuses on the pair-wise strong interference links. Up to now, the adopted hypergraph models in existing studies for wireless communications are belonging to the undirected hypergraph category. However, undirected hypergraph interference model cannot precisely reflect the asymmetric interference. For a undirected hyperedge, there exists at least one node interfered by others' co-channel transmission, but we cannot clearly classify the interfered nodes and non-interfered nodes and there may exist multiple possible interference patterns as shown in Fig. 2 (more details will be given in the following). To capture the asymmetric interference effect in hyperedges, we resort to the directed hypergraph and the corresponding definition is given as follows:

Definition 3 ((Directed hypergraph) [10]): A directed hypergraph is a hypergraph with directed hyperedges. A directed hyperedge or hyperarc is an ordered pair, $E = (X, Y)$, of (possibly empty) disjoint subsets of vertices; X is the tail of E while Y is its head. In the following, the tail and the head of hyperarc E will be denoted by $T(E)$ and $H(E)$, respectively.

As for directed graphs, the incidence matrix of a hypergraph H an $n \times m$ matrix $A = [a_{ij}]_{n \times m}$ defined as follows:

$$a_{ij} = \begin{cases} 1, & \text{if } v_i \in H(E_j), \\ -1, & \text{if } v_i \in T(E_j), \\ 0, & \text{otherwise.} \end{cases} \tag{1}$$

B. Directed Hypergraph Model With Asymmetric Interference

Generally, there is a one-to-one correspondence between $(-1, 0, 1)$ incidence matrix and hypergraph. In Fig. 3, we

---

[1]Since there exist numerous studies on the construction of hypergraph interference model as in [1], in this paper, we only focus on channel allocation with given directed hypergraph interference model.
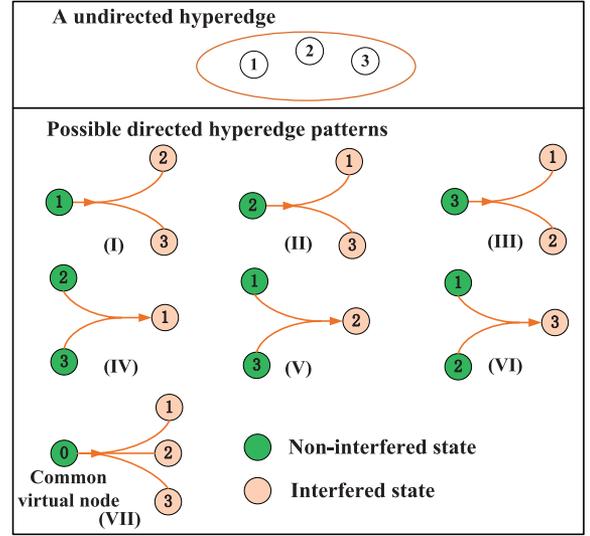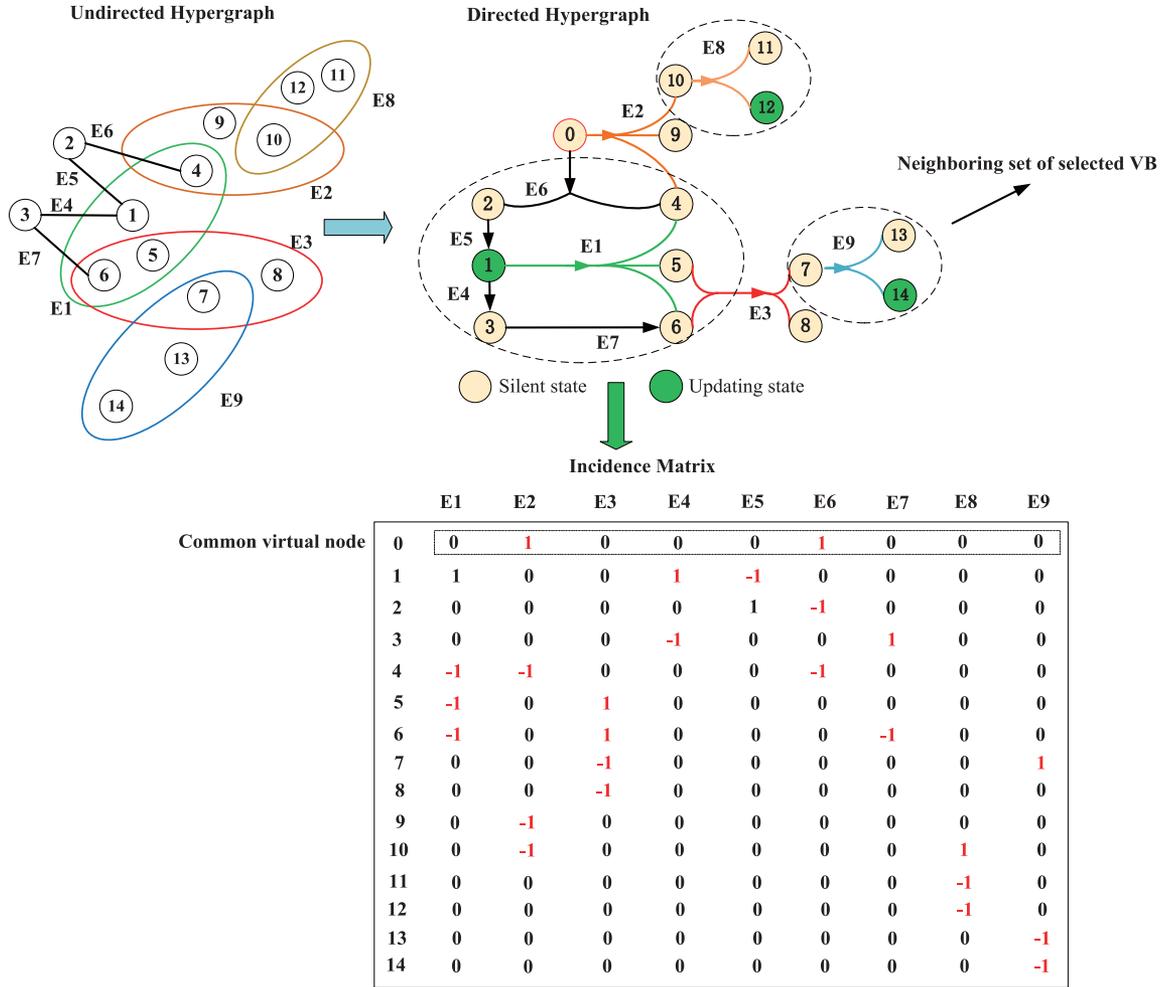
Fig. 3.   Illustration of the improved directed hypergraph interference model and its incidence matrix. Note that in undirected hypergraph, hyperedges of cardinality 2 are represented as regular binary edges and the hyperedges of larger cardinality are represented as circles containing a set of vertices.

present an illustration of a directed hypergraph, consisting of 14 vertices and 9 corresponding hyperedges (marked by E1-E9 for simplicity), and its corresponding incidence matrix. We use interchangeably 'hyperedge' with 'hyperarc' for directed hypergraph. For hyperedge E1, node 1 is in its head set $H(E1)$ and nodes 4, 5 and 6 constitute their tail set $T(E1)$. To apply the directed hypergraph for interference relationship model in wireless networks, we need to do slight adjustments for the original directed hypergraph. For some special undirected hyperedges, such as E2, in communication network, its nodes are all interfered by their co-channel transmission, then we generate a *common virtual node 0* in the head set of the corresponding hyperarcs, avoiding the head set to be empty and meaningless. In Fig. 2, we show the possible directed hyperedge patterns for a given undirected hyperedge with $Q = 3$, where Q is the maximum cardinality of hyperedge, in wireless communications. Vertices in each directed hyperedge can be marked with one of two different states ('−1' and '1'). This motivates us to utilize the two states to represent nodes' interference states ('interfered' or 'non-interfered') in the corresponding hyperedges as shown in Fig. 3, when co-channel transmission occurs. Note that the

state of the common virtual node 0 is always marked with '1' (i.e., 'non-interfered').

C. Optimization Problem

In the following, we first introduce a binary variable $s_i(e_i^k)$, denoting node i's interference state in hyperedge e_i^k to reflect the asymmetric interference, as follows:

$$s_i(e_i^k) = \begin{cases} 1; & \text{if } a_{i,e_i^k} = -1; \\ 0; & \text{else.} \end{cases} \quad (2)$$

Definition 2 (Protocol interference, PI): Let $I_i(a_i)$ be node i's received cumulate protocol interference from its adjacent hyperedges when accessing channel a_i:

$$I_i(a_i, a_{-i}) = \sum_{e_i^k \in \mathbf{E}(i)} \delta\left(e_i^k\right) s_i(e_i^k), \quad (3)$$

where $a_i \in \mathbf{A}_i$ and a_{-i} denote node i's selected channel and the channel selection profile of all nodes except node i, respectively. $\mathbf{A}_i$ is node i's available channel set. For simplicity, assume $\mathbf{A}_i = \{0, 1, 2, \dots, M\}$ for any node i in the directed hypergraph and

we additionally introduce a *virtual channel 0* which represents sleeping strategy. Moreover, $\mathbf{E}(i) = \left[e_i^1, e_i^2, \ldots, e_i^{|\mathbf{E}(i)|} \right]$ represents the hyperedge set of vertex i and note that each hyperedge in it contains the vertex i; $| \bullet |$ represents the cardinality of the set; indicator function $\delta (\bullet)$ is defined as follows:

$$\delta \left(e_i^k \right) = \begin{cases} 1; & a_i \neq 0; a_j = a_i; \forall j \in e_i^k; \\ 0; & \text{otherwise.} \end{cases} \quad (4)$$

Let C_i denote the normalized capacity of node i as

$$C_i(a_i, a_{-i}) = \begin{cases} 1; & I_i(a_i, a_{-i}) = 0, a_i \neq 0 \\ 0; & \text{otherwise,} \end{cases} \quad (5)$$

where $C_i(a_i, a_{-i}) = 0$ implies that node i cannot correctly decode the packet or node i is in sleeping state. Define the normalized network capacity $C(a_1, a_2, \ldots, a_N)$ as[2]

$$C(a_1, a_2, \ldots, a_N) = \sum_{i \in N} C_i(a_i, a_{-i}). \quad (6)$$

Hence, we construct the optimization problem as follows:

$$\max_{\mathbf{a} = (a_1, a_2, \ldots, a_N)} \quad C(\mathbf{a})$$

$$\text{s.t.} \qquad a_i \in \mathbf{A}_i, i = 1, 2, \ldots, N. \quad (7)$$

The problem in (7) is a combinatorial optimization problem, which is proved to be NP-hard [1].

D. Game Model and Equilibrium Analysis

We formulate a channel access game $G = \left[\mathbf{N}, \left\{ \mathbf{A}_i \right\}_{i \in \mathbf{N}}, \left\{ U_i \right\}_{i \in \mathbf{N}} \right]$ for VBs in cloud, where U_i is player i's utility function. Recall that there exists a one-to-one correspondence between VBs and physical communication nodes. Inspired by [9], we define player (VB) i's utility function in a locally altruistic form as

$$U_i(a_i, a_{-i}) = C_i(a_i, \mathbf{a}_{N_i}) + \sum_{k \in N_i} C_k(a_k, \mathbf{a}_{N_k}) \chi_{i,k}, \quad (8)$$

where N_i denotes player i's neighboring set, including those VBs except VB 0 connecting VB i via directed hyperedges; $\mathbf{a}_{N_i}$ denotes the channel selection profile of VB i's neighbors. $\chi_{i,k}$ denotes the binary correlation factor as

$$\chi_{i,k} = \begin{cases} 0; & \forall e_k^m \in \mathbf{E}(i, k), s_k(e_k^m) = 0; \\ 1; & \text{otherwise.} \end{cases} \quad (9)$$

Note that, $\chi_{i,k} = 0$ implies node i is not interfered by the neighboring node k's transmission behavior and we term the VB k as a "uncoupled neighbor" of VB i. For instance, in Fig. 2, both VB 7 and 8 are "uncoupled neighbors" of VB 5.

Each player aims to maximize its utility as follows:

$$\max_{a_i \in \mathbf{A}_i} U_i(a_i, a_{-i}), \forall i \in \mathbf{N}. \quad (10)$$

We present two important definitions in the following.

---

[2]We can adopt other performance metrics, such as network throughput, quality of experience and transmission delay, for evaluation. The proposed learning algorithm can work with a slight change in utility function structure.

Definition 3 (Pure Nash equilibrium): An action profile $\mathbf{a}^* = (a_1^*, a_2^*, \ldots, a_N^*)$ is PNE if no player can obtain more profit by deviating unilaterally. Mathematically, the PNE satisfies $U_i(a_i^*, a_{-i}^*) \geq U_i(a_i, a_{-i}^*), \forall i \in \mathbf{N}, \forall a_i \in \mathbf{A}_i, a_i \neq a_i^*$.

Definition 4 (Exact potential game): A game G is an EPG if there exists an exact potential function $\phi : \mathbf{A}_1 \times \mathbf{A}_2 \times \cdots \mathbf{A}_N \to \mathbb{R}$, which satisfies $U_i(a_i, a_{-i}) - U_i(\overline{a}_i, a_{-i}) = \phi(a_i, a_{-i}) - \phi(\overline{a}_i, a_{-i})$, for any two actions $a_i \in \mathbf{A}_i$ and $\overline{a}_i \in \mathbf{A}_i$ of any player i.

For an EPG, it possesses at least one PNE which maximizes the potential function [9].

Theorem 1: The proposed game G is an EPG.

Proof: We construct the potential function as $\Phi(a_i, a_{-i}) = \sum_{i \in \mathbf{N}} C_i(a_i, a_{-i}) = C(\mathbf{a})$. Now, assume that an arbitrary player i unilaterally changes its action from a_i to $\overline{a}_i$, the change in its utility function is:

$$U_i(\overline{a}_i) - U_i(a_i) = [C_i(\overline{a}_i, \mathbf{a}_{N_i}) - C_i(a_i, \mathbf{a}_{N_i})]$$
$$+ \left[\sum_{k \in N_i} (C_k(a_k, \overline{\mathbf{a}}_{N_k}) \chi_{i,k} - C_k(a_k, \mathbf{a}_{N_k}) \chi_{i,k}) \right]. \quad (11)$$

The change of the potential function corresponding to player i's unilateral deviation is given by

$$\Phi(a_i, a_{-i}) - \Phi(\overline{a}_i, a_{-i}) = \sum_{i \in \mathbf{N}} C_i(a_i, a_{-i}) - \sum_{i \in \mathbf{N}} C_i(\overline{a}, a_{-i})$$

$$= [C_i(a_i, \mathbf{a}_{N_i}) - C_i(\overline{a}_i, \mathbf{a}_{N_i})]$$

$$+ \left[\sum_{k \in N_i} (C_k(a_k, \mathbf{a}_{N_k}) - C_k(a_k, \overline{\mathbf{a}}_{N_k})) \right]$$

$$+ \left[\sum_{l \notin N_i, l \neq i} (C_l(a_l, \mathbf{a}_{N_l}) - C_l(a_l, \overline{\mathbf{a}}_{N_l})) \right]. \quad (12)$$

Since $\sum_{l \notin N_i, l \neq i} (C_l(a_l, \mathbf{a}_{N_l}) - C_l(a_l, \overline{\mathbf{a}}_{N_l})) = 0$, and note that

$$\sum_{k \in N_i} (C_k(a_k, \mathbf{a}_{N_k}) - C_k(a_k, \overline{\mathbf{a}}_{N_k}))$$

$$= \sum_{k \in N_i} (C_k(a_k, \overline{\mathbf{a}}_{N_k}) \chi_{i,k} - C_k(a_k, \mathbf{a}_{N_k}) \chi_{i,k})$$

$$+ \sum_{k \in N_i} (C_k(a_k, \overline{\mathbf{a}}_{N_k})(1 - \chi_{i,k}) - C_k(a_k, \mathbf{a}_{N_k})(1 - \chi_{i,k})). \quad (13)$$

Next, we check that, for any $k \in N_i$, the following equation holds:

$$C_k(a_k, \overline{\mathbf{a}}_{N_k})(1 - \chi_{i,k}) - C_k(a_k, \mathbf{a}_{N_k})(1 - \chi_{i,k}) = 0. \quad (14)$$

For any node k satisfying $\chi_{i,k} = 1$, (14) holds obviously; on the other hand, for any node k satisfying $\chi_{i,k} = 0$, node k's interference state in hyperedges is not affected by node i, hence, we can get $C_k(a_k, \overline{\mathbf{a}}_{N_k}) = C_k(a_k, \mathbf{a}_{N_k})$, implying (14) still holds.

Hence, (13) can be simplified as follows:

$$\sum_{k \in N_i} \left(C_k(a_k, \mathbf{a}_{N_k}) - C_k(a_k, \overline{\mathbf{a}}_{N_k}) \right)$$

$$= \sum_{k \in N_i} \left(C_k(a_k, \overline{\mathbf{a}}_{N_k}) \chi_{i,k} - C_k(a_k, \mathbf{a}_{N_k}) \chi_{i,k} \right). \qquad (15)$$

Therefore, we can verify that the following equation holds:

$$\Phi(\overline{a}_i, a_{-i}) - \Phi(a_i, a_{-i}) = U_i(\overline{a}_i) - U_i(a_i). \qquad (16)$$

That completes the proof. ∎

Thus, the optimal PNE of the proposed game G corresponds to the optimal solution of (7).

III. CENTRALIZED-DISTRIBUTED LEARNING IN CLOUD

A number of learning algorithms are available to facilitate the convergence of potential games to NE, e.g., best-response dynamic, spatial adaptive play (SAP), and fictitious play. While, not all of these algorithms can guarantee to achieve the optimal PNE. Inspired by SAP, we design a multi-agent concurrent learning algorithm (MACL), which runs in a centralized-distributed fashion, to find the optimal PNE of the formulated game. To accelerate the convergence process, MACL randomly selects multiple concurrent VBs in directed hypergraph to update actions rather than one in SAP at each iteration. For any two selected VBs i and j, $N_i \cap N_j = \emptyset$, i.e., the intersection of any two selected VBs' neighboring set is empty. Hence, the selected VBs are located in separated interference region, and any selected VB's action change will not affect others' utility. Specifically, in step 4, the MACL first randomly selects one updating VB and then marks his neighbors and neighbors' neighbors as non-updating state. For those remaining unmarked nodes, MACL successively implements the rule mentioned above until all VBs' states are determined. For instance, VB 1, 12 and 14 in Fig. 3 are selected to update their actions at one iteration step, and we mark the selected VBs' neighboring sets in the directed hypergraph.

In the k-th iteration, let $\Lambda_{\mathbf{A}_i}(k)$ represent the VB i's action probability distribution over action set $\mathbf{A}_i$ and $q_i^{a_i}(k) \in \Lambda_{\mathbf{A}_i}(k)$ denotes the selected probability of action $a_i(k)$. The selected VBs update their actions with probability following the Boltzmann-Gibbs rule as:

$$q_i^{a_i}(k+1) = \frac{\exp\left(\beta U_i(a_i, \mathbf{a}_{N_i}(k))\right)}{\sum_{\overline{a}_i \in \mathbf{A}_i} \exp\left(\beta U_i(\overline{a}_i, \mathbf{a}_{N_i}(k))\right)}, \qquad (17)$$

where β is the learning parameter which satisfies $\beta > 0$ and can be interpreted as the rational level of the players. Furthermore, MACL is implemented in the cloud and the information exchange among VBs can be easily done with a limited cost. Until MACL converges, VBs feedback the final result to the physical devices. The details of MACL are given in Algorithm 1.

Theorem 2: If the learning parameter β is sufficiently large, MACL can maximize the normalized network capacity with an arbitrarily high probability.

Proof: We will give a sketch proof following the similar lines in [9]. In the Theorem 4 of [9], it is proved that SAP can achieve the global maxima of the potential function with an arbitrarily

---

Algorithm 1: Centralized-Distributed Multi-Agent Concurrent Learning Algorithm (MACL).

Step 1: WSP maps the physical communication networks to a VDN in the cloud.

Step 2: Collect the context information of each node and construct the directed hypergraph interference model.

Step 3: All the VBs (players) exchange information with their 'coupled' neighbors.

Step 4: In the k-th iteration, randomly select multiple concurrent VBs in hypergraph. For any two selected VBs i and j, $N_i \cap N_j = \emptyset$. The selected brokers update their actions probability distribution $\Lambda_{\mathbf{A}_i}(k)$ as (19). Then, the selected players randomly select an action according to the $\Lambda_{\mathbf{A}_i}(k+1)$. Other players' actions maintain unchanged as $a_i(k+1) = a_i(k)$.

Step 5: If the predefined maximum number of iteration steps $k_{\max}$ is reached, stop; else go to Step 2.

Step 6: VBs feedback the learning result to the corresponding physical devices.

---

high probability. In MACL, the randomly selected VBs at each iteration do not affect each other in the improved directed interference hypergraph. MACL can be viewed as the concurrent version of SAP for the improved directed hypergraph. Therefore, MACL can converge to the global optimum of potential function and maximize the normalized network capacity with an arbitrary high probability. ∎

There is no accurate complexity analysis for learning algorithm. Roughly speaking, at each iteration of SAP, one player is randomly chosen to alter its action after calculating $M+1$ available rewards from its channel set. Accordingly, the complexity of SAP is $O(T(M+1))$, where T is the number of iteration steps. For the proposed MACL, at one iteration, the total complexity is $O(N(M+1))$, because the complexity of multiple VBs' selection is $O(N)$. Similarly, the overall complexity of MACL is $O(TN(M+1))$. Note that convergence rate of MACL is faster than SAP, so MACL can converge with significantly less iteration steps. Furthermore, the complexity burden of each selected VB is only $O(M+1)$ at one iteration. For the hypergraph-coloring scheme in [1], the computational complexity is $O(N^3)$.

On the other hand, MACL scheme tends to be more robust to network topology dynamic, because the changes in topology may only affect the devices in a local region, and their corresponding VBs can quickly learn the new access strategy with a relatively low computation complexity.

IV. SIMULATION RESULTS

In this section, simulations are conducted to evaluate the performance of the proposed MACL. We construct random directed hypergraph with $Q = 3$ to model the ultra-dense communication
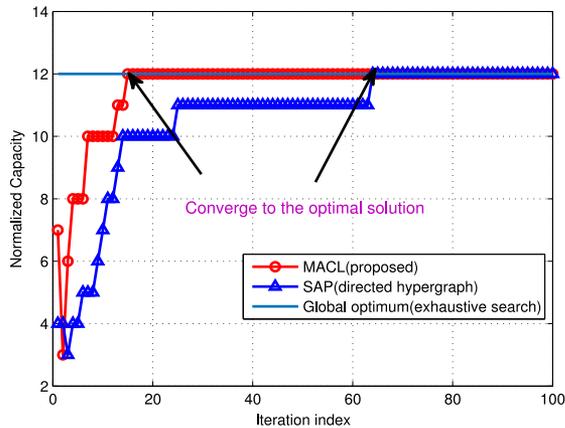
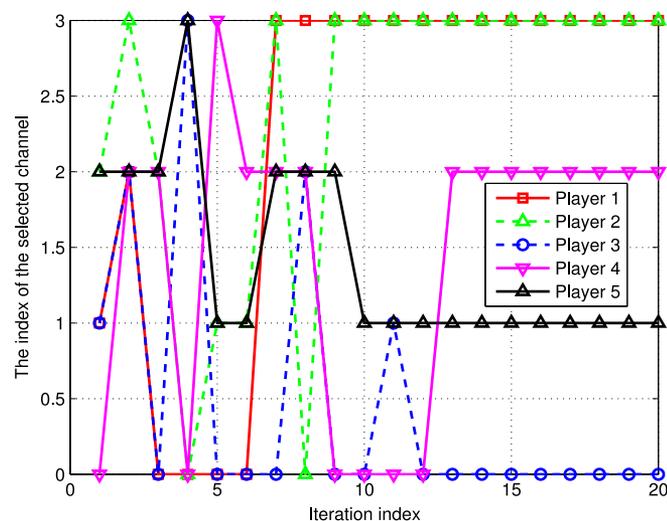Fig. 4. Convergence behavior of the proposed learning algorithm.

Fig. 5. Partial players' strategy update in the learning process of MACL.
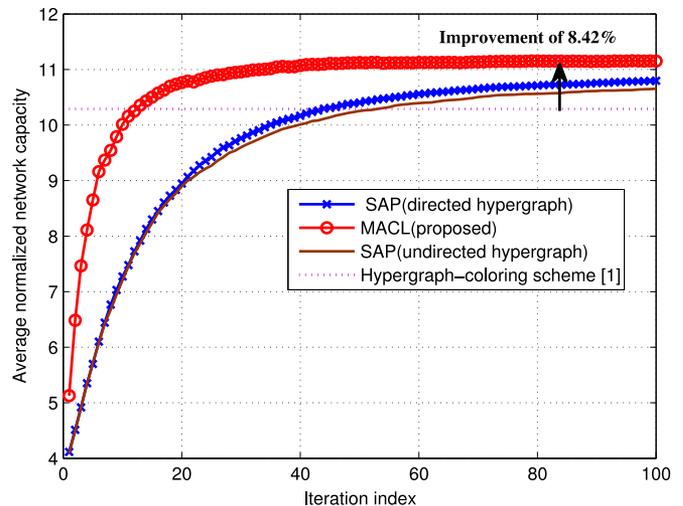

Fig. 6. Average normalized network capacity versus iteration steps.

the proposed MACL in Fig. 6. In the current system configuration, MACL can get an improvement of 8.42% compared with hypergraph-coloring scheme [1]. The reason lies in that MACL fully considers the asymmetric interference relationship; hence, the channel allocation strategy is more appropriate and close to the optimal solution. On the other hand, the performance of MACL is better than SAP (for both directed and undirected hypergraph) because of its fast convergence rate within a given $k_{\max}$.

V. CONCLUSION

In this paper, we studied the channel allocation for ultra-dense D2D communications based on directed hypergraph interference model. We formulated this problem as a local cooperation game in the cloud, which is proved to be an exact potential game admitting at least one PNE. Then, we designed a multi-agent concurrent learning algorithm to converge to the optimal PNE, which yields an optimal solution for the given optimization problem.

scenario in a relatively small region.[3] The number of available channels M and active end users N are set as 3 and 13, respectively. We set the learning parameter $\beta = k$, which is increasing with iteration index k and bounded by 50. Maximum number of iteration steps is set as $k_{\max} = 100$. In Fig. 4, we present the convergence performance comparison of MACL with SAP for a random directed hypergraph. We can observe that both MACL and SAP can achieve the optimal solution, which is obtained by exhaustive search. Moreover, MACL converges significantly faster than SAP.

In Fig. 5, we present partial players' strategy update in the learning process of MACL. It is seen that multiple players can update their actions simultaneously in the same iteration; for instance, at the 6-th iteration, player 1, 2 and 5 are selected to update their strategies. Note that player 3 finally selects channel 0, which stands for the sleeping strategy.

We generate 1000 random D2D networks, and then evaluate the performance metric of the average normalized capacity for

---

[3]As shown in [1], [5], $Q = 3$ can achieve a good tradeoff between performance approximation and computation complexity for hypergraph.

REFERENCES

[1] H. Zhang, L. Song, and Z. Han, "Radio resource allocation for device-to-device underlay communications using hypergraph theory," *IEEE Trans. Wireless Commun.*, vol. 15, no. 7, pp. 4852–4861, Jul. 2016.

[2] Y. Sun, Q. Wu, Y. Xu, Y. Zhang, F. Sun, and J. Wang, "Distributed channel access for device-to-device communications: A hypergraph-based learning solution," *IEEE Commun. Lett.*, vol. 21, no. 1, pp. 180–183, Jan. 2017.

[3] S. Sarkar and K. N. Sivarajan, "Hypergraph models for cellular mobile communication systems," *IEEE Trans. Veh. Technol.*, vol. 47, no. 2, pp. 460–471, May 1998.

[4] J. Feng and M. Tao, "Hypergraph-based frequency reuse in dense femtocell networks," in *Proc. IEEE Int. Conf. Commun. China*, Xi'an, China, Aug. 2013, pp. 537–542.

[5] Q. Li and R. Negi, "Maximal scheduling in wireless ad hoc networks with hypergraph interference models," *IEEE Trans. Veh. Technol.*, vol. 61, no. 1, pp. 297–310, Jan. 2012.

[6] Z. Du, Y. Sun, W. Guo, Y. Xu, Q. Wu, and J. Zhang, "Data-driven deployment and cooperative self-organization in ultra-dense small cell networks," *IEEE Access*, vol. 6, pp. 22839–22848, 2018.

[7] Q. Wu, G. Ding, Z. Du, Y. Sun, M. Jo, and A. V. Vasilakos, "A cloud-based architecture for internet of spectrum devices (IoSDs) over 5G networks," *IEEE Access*, vol. 4, pp. 2854–2862, Jun. 2016.

[8] Y. Sun, Q. Wu, J. Wang, Y. Xu, and A. Anpalagan, "VERACITY: Overlapping coalition formation based double auction for heterogeneous demand and spectrum reusability," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 10, pp. 1–16, Oct. 2016.

[9] Y. Xu, J. Wang, Q. Wu, Anpalgan, and Y.-D. Yao, "Opportunistic spectrum access in cognitive radio networks: Global optimization using local interaction games," *IEEE J. Sel. Topics Signal Process.*, vol. 6, no. 2, pp. 180–194, Apr. 2012.

[10] G. Gallo, G. Longo, and S. Pallottinon, "Directed hypergraphs and applications," *Discr. Appl. Math.*, vol. 42, no. 2–3, pp. 177–201, 1993.

[11] A. Bretto, *Hypergraph Theory: An Introduction*. Berlin, Germany: Springer, 2013.

[12] S. Lasaulce and H Tembine, *Game Theory and Learning for Wireless Networks*. Amsterdam, The Netherlands: Elsevier, 2011.

[13] H. Zhang, L. Song, Y. Li, and G. Y. Li, "Hypergraph theory: Applications in 5G heterogeneous ultra-dense networks," *IEEE Commun. Mag.*, vol. 55, no. 12, pp. 70–76, Dec. 2017.

Youming Sun received the B.S. degree in electronic and information engineering from Yanshan University, Qinhuangdao, China, in 2010 and the Ph.D. degree from National Digital Switching System Engineering and Technological Research Center, Zhengzhou, China, in 2016, respectively. His research interests include resource allocation in small cell networks, cognitive radio networks, UAV communication networks, game theory, and statistical learning.

He currently serves as a regular reviewer for many technical journals, including the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS, the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, the IEEE SYSTEMS JOURNAL, the IEEE Access, *Transactions on Emerging Telecommunications Technologies*. He has acted as Technical Program Committees member for IEEE International Conference on Wireless Communications and Signal Processing 2015.
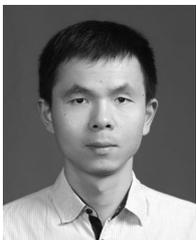
Zhiyong Du (S'12–M'17) received the B.S. degree in electronic information engineering from the Wuhan University of Technology, Wuhan, China, in 2009 and the Ph.D. degree in communications and information systems from the College of Communications Engineering, PLAUST, Nanjing, China, in 2015. Since 2016, he has been an Assistant Professor with the National University of Defense Technology, Changsha, China. His research interests include 5G, quality of experience, learning theory, and game theory.
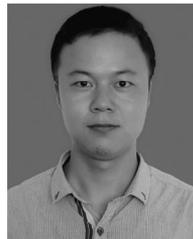
Yuhua Xu (S'08–M'13) received the B.S. degree in communications engineering and the Ph.D. degree in communications and information systems from the College of Communications Engineering, PLA University of Science and Technology, Nanjing, China, in 2006 and 2014, respectively. He is currently an Associate Professor with the College of Communications Engineering, Army Engineering University of PLA. He has authored several papers in international conferences and reputed journals in his research area. His research interests focus on UAV communication networks, opportunistic spectrum access, learning theory, and distributed optimization techniques for wireless communications.

Dr. Xu was the recipient of the IEEE Signal Processing Society 2015 Young Author Best Paper Award and the Funds for Distinguished Young Scholars of Jiangsu Province in 2016. He received the Certificate of Appreciation as an Exemplary Reviewer for the IEEE COMMUNICATIONS LETTERS in 2011 and 2012, respectively. He served as an Associate Editor for the Transactions on *Emerging Telecommunications Technologies* (Wiley) and the *KSII Transactions on Internet and Information Systems* and a Guest Editor of Special Issue on The Evolution and the Revolution of 5G Wireless Communication Systems for the *IET Communications*.

Yuli Zhang (S'16) received the B.S. degree in electronic and information engineering from Peking University, Beijing, China, in 2012, and the M.S. degree from the College of Communications Engineering, Army Engineering University of PLA, Nanjing, China, in 2015, respectively, where he is currently working toward the Ph.D. degree. His research interests include resource allocation in small cell networks, game theory, and energy harvesting.

Luliang Jia (S'17) received the B.S. degree in communications engineering from Lanzhou Jiaotong University, Lanzhou, China, in 2011 and the M.S. degree in communications and information systems from College of Communication Engineering, Army Engineering University of PLA, Nanjing, China, in 2014, respectively. He is currently working toward the Ph.D. degree from the College of Communication Engineering, Army Engineering University of PLA. His current research interests include game theory, learning theory, and communication anti-jamming technology.

Alagan Anpalagan (S'98–M'01–SM'04) received the BASc., MASc. and Ph.D. degrees in electrical engineering from the University of Toronto, Toronto, ON, Canada. He is currently a Professor with the Department of Electrical and Computer Engineering, Ryerson University, Toronto, ON, Canada, where he directs a research group working on radio resource management and radio access and networking areas within the WINCORE Lab. He is a registered Professional Engineer in the province of Ontario, Canada. He was an Editor of *EURASIP Journal of Wireless Communications and Networking* (2004–2009), IEEE COMMUNICATION LETTER *(2010–2013), Springer Wireless Personal Communications (2011–2013)* (2011–2013), ands the IEEE COMMUNICATIONS SURVEYS AND TUTORIALS (2012–2014). He coauthored three edited books, *Design and Deployment of Small Cell Networks*, Cambridge Univ. Press (2014), *Routing in Opportunistic Networks*, Springer (2013), *Handbook on Green Information and Communication Systems*, Academic (2012).

Dr. Anpalagan is currently serves as TPC Vice-Chair, IEEE VTC Fall-2017, and served as TPC Co-Chair, IEEE Globecom15: SAC Green Communication and Computing, IEEE WPMC'2012 Wireless Networks, IEEE PIMRC'2011 Cognitive Radio and Spectrum Management, and IEEE CCECE'2004/2008. He served as Vice Chair, IEEE SIG on Green and Sustainable Networking and Computing with Cognition and Cooperation (2015), IEEE Canada Central Area Chair (2012–2014), IEEE Toronto Section Chair (2006–2007), ComSoc Toronto Chapter Chair (2004–05), and IEEE Canada Professional Activities Committee Chair (2009–2011). He is the recipient of the Deans Teaching Award (2011), Faculty Scholastic, Research and Creativity Award (2010, 2014, and 2017), Faculty Service Award (2011 and 2013) in Ryerson University and also the recipient of Exemplary Editor Award from IEEE ComSoc (2013) and Editor-in-Chief Top 10 Choice Award in Transactions on Emerging Telecommunications Technology (2012). He is a Fellow of the Institution of Engineering and Technology.